REGULAR LAGRANGE MULTIPLIERS FOR CONTROL PROBLEMS WITH MIXED POINTWISE CONTROL-STATE CONSTRAINTS

FREDI TRÖLTZSCH[1]

Abstract. A class of quadratic optimization problems in Hilbert spaces is considered, where pointwise box constraints and constraints of bottleneck type are given. The main focus is to prove the existence of regular Lagrange multipliers in L^2-spaces. This question is solved by investigating the solvability of a Lagrange dual quadratic problem. The theory is applied to different optimal control problems for elliptic and parabolic partial differential equations with mixed pointwise control-state constraints.

Key words. Quadratic programming, Hilbert space, regular Lagrange multipliers, optimal control, elliptic and parabolic equation, mixed pointwise control-state constraint.

AMS subject classifications. 49K20, 49N10, 49N15, 90C45

1. A class of optimization problems. We consider the following general class of optimization problems in Hilbert spaces

$$(PP) \qquad \text{minimize} \quad \frac{1}{2}\,\|S\,u\|_H^2 + \int_D (a(x)\,u(x) + \frac{\kappa}{2}\,u(x)^2)\,dx$$

subject to the pointwise constraints

$$\begin{aligned} u(x) &\le c(x) + (G\,u)(x) \\ u(x) &\le b(x) \\ u(x) &\ge 0 \end{aligned}$$

to be fulfilled for almost all $x \in D$, where the function u is taken from $L^2(D)$.

In this problem, henceforth called *primal problem*, $D \subset \mathbb{R}^N$ is a Lebesgue-measurable bounded set, S and G are linear continuous operators from $L^2(D)$ to a real Hilbert space H and $L^2(D)$, respectively, while a, b, c are fixed functions from $L^2(D)$ and $\kappa > 0$ is a fixed constant. Let us denote the natural inner product of $L^2(D)$ by $(\cdot\,,\,\cdot)$ and the associated norm by $\|\cdot\|$. In all what follows, we denote the primal objective function by f, i.e.

$$f(u) = \frac{1}{2}\|S\,u\|_H^2 + \frac{\kappa}{2}\,\|u\|^2 + (a\,,\,u).$$

A function $u \in L^2(D)$ is said to be *feasible*, if it satisfies the pointwise constraints given above.

This class of problems is sufficiently large to cover several types of linear-quadratic optimal control problems for elliptic or parabolic equations with pointwise control-state constraints. Below, we briefly sketch some possible examples. In all of these problems, the regularity of Lagrange multipliers associated with the state constraints is an important issue. The main aim of our paper is to show that, under natural assumptions, the multipliers can assumed to be functions from $L^2(D)$.

Examples. The following optimal control problems for partial differential equations are covered by (PP).

[1]Institut für Mathematik, Technische Universität Berlin, D-10623 Berlin, Str. des 17. Juni 136, Germany.

(i) *Elliptic control problem with distributed control and distributed observation*

$$\min \frac{1}{2}\,\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\kappa}{2}\,\|u\|^2_{L^2(\Omega)}$$

subject to

$$\begin{array}{rlrl} -\Delta\, y(x) = u(x) & \text{in } \Omega & 0 \le u(x) \le b(x) & \text{in } \Omega \\ y(x) = 0 & \text{on } \Gamma & u(x) - y(x) \le c(x) & \text{in } \Omega, \end{array}$$

where $\Omega \subset \mathbb{R}^N$, $N \ge 2$, is a bounded domain with $C^{0,1}$-boundary Γ and y_Ω is a fixed function from $L^2(\Omega)$. The function u is the *control* while y is the associated *state*. The constraints $0 \le u \le b$ are called *box constraints* while the restriction $u - y \le c$ is said to be a *mixed control state constraint*.

It is known that the elliptic equation admits for each $u \in L^2(\Omega)$ a unique weak solution $y \in H^1_0(\Omega)$. The solution mapping $\hat{S} : u \mapsto y$ is continuous. Let us denote by E the embedding operator from $H^1_0(\Omega)$ to $L^2(\Omega)$

This problem is related to (PP) as follows: We define S as the solution operator, considered with range in $H = L^2(\Omega)$, that is $S = E\,\hat{S}$, $S : L^2(\Omega) \to L^2(\Omega)$. Expanding the first norm square, we get $\frac{1}{2}\,\|y - y_\Omega\|^2_{L^2(\Omega)} = \frac{1}{2}\,\|y\|^2_{L^2(\Omega)} - (S^\star y_\Omega\,,\,u) +$ const. , hence $a = -\,S^\star y_\Omega$. The operator G is defined by $G = S$, i.e. the pointwise control-state constraints are considered in $L^2(D)$, hence $D = \Omega$.

(ii) *Elliptic boundary control problem with distributed observation*

We consider a similar case of boundary control and introduce some more coefficients to make the setting more flexible for applications.

$$\min \frac{1}{2}\,\|\alpha\, y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\kappa}{2}\,\|u\|^2_{L^2(\Gamma)}$$

subject to

$$\begin{array}{rlrl} -\Delta\, y(x) = 0 & \text{in } \Omega & 0 \le u(x) \le b(x) & \text{on } \Gamma \\ \partial_n\, y(x) + \beta(x)\, y(x) = u(x) & \text{on } \Gamma & u(x) - \gamma(x)\, y(x) \le c(x) & \text{on } \Gamma, \end{array}$$

where we use the same quantities as in (i), and $\alpha \in L^\infty(\Omega)$, β, $\gamma \in L^\infty(\Gamma)$ are fixed non-negative functions. The solution mapping $\hat{S} : u \mapsto y$ is continuous from $L^2(\Gamma)$ to $H^1(\Omega)$. Let E be the embedding operator of $H^1(\Omega)$ into $L^2(\Omega)$. Now, $D = \Gamma$, $H = L^2(\Omega)$, $S = \alpha\, E\, \hat{S}$, $S : L^2(\Gamma) \to L^2(\Omega)$, and $G = \gamma\, \tau\, \hat{S} : L^2(\Gamma) \to L^2(\Gamma)$, where $\tau : H^1(\Omega) \to L^2(\Gamma)$ denotes the trace operator. By ∂_n the normal derivative with respect to the outward normal vector is denoted.

(iii) *Parabolic boundary control problem*

Let us mention a further application:

$$\min \frac{1}{2}\,\|\alpha\, y(T) - y_\Omega\|^2_{L^2(\Omega)} + \frac{\kappa}{2}\,\|u\|^2_{L^2(\Sigma)}$$

subject to

$$\begin{array}{rl} y_t(x,t) - \Delta\, y(x,t) = 0 & \text{in } Q \\ \partial_n\, y(x,t) + \beta(x,t)\, y(x,t) = u(x,t) & \text{on } \Sigma \\ y(x,0) = 0 & \text{in } \Omega, \end{array}$$

$$\begin{aligned} 0 \le u(x,t) &\le b(x,t) && \text{on } \Sigma \\ u(x,t) - \gamma(x,t)\, y(x,t) &\le c(x,t) && \text{on } \Sigma. \end{aligned}$$

Here, the notations $Q = \Omega \times (0,T)$, $\Sigma = \Gamma \times (0,T)$ are used. Moreover, functions $\alpha \in L^\infty(\Omega)$, $\beta \in L^\infty(\Gamma)$, $\gamma \ge 0 \in L^\infty(\Sigma)$ are given. It is known that the solution mapping $\hat{S} : u \mapsto y$ is continuous from $L^2(\Sigma)$ to

$$W(0,T) = \{ y \in L^2(0,T;H^1(\Omega)) \,|\, y_t \in L^2(0,T;H^1(\Omega)') \}.$$

Therefore, the mapping $S : u \mapsto \alpha\, y(T)$ is well defined and continuous from $L^2(\Sigma)$ to $H = L^2(\Omega)$, and $G : u \mapsto \gamma\, \tau y$ is continuous in $L^2(\Sigma)$. The choice $D = \Gamma$ fits into our setting.

In all of these examples, the operator G is non-negative, i.e. it transforms non-negative functions u into non-negative functions. This follows from the maximum principle for elliptic and parabolic PDEs.

Problems with pointwise control-state constraints of the type defined above have already been the subject of various papers devoted to *bottleneck problems*. In the sixties, they were considered extensively for linear programming problems in Banach spaces in the context of optimal control of ODEs, see [6], [11]. In [10], the author extended these ideas to the case of parabolic equations to derive necessary optimality conditions in form of a minimum principle.

Later, more general linear and semilinear parabolic equations were discussed [1], [2], [3], and second-order sufficient optimality conditions were derived for the nonlinear parabolic case in [9]. Moreover, the Pontryagin maximum principle was proven for parabolic problems with mixed control-state constraints in [4].

In early papers on this subject, techniques of linear programming in Banach spaces were used. The Lagrange multiplier for the state constraints was obtained as the solution of a dual linear problem. The main difficulty was the proof of existence of solutions to the dual problem. This technique can also be extended to nonlinear problems by showing that an optimal control solves a linearized problem. Then the Lagrange multiplier can be obtained as solution of the dual to the linearized program. This technique was applied by Bergounioux and the author in [2], [3]. Later, Arada and Raymond [1] introduced a more direct approach starting from the existence of Lagrange multipliers in the space $L^\infty(D)^\star$. This existence is directly obtained from the Karush-Kuhn-Tucker theory in Banach spaces. Then the crucial step is to show higher regularity by a compactification technique, a quite general but rather technical approach.

In [2], [3] the inverse-monotonicity of the operator $(I - G)^{-1}$ is employed, which is only true, if the associated Neumann series $I + G + G^2 ...$ converges. In the elliptic case, that was the starting point of this paper, this convergence only holds true for a "small" operator G. In the example (ii), this means that γ must have a sufficiently small L^∞-norm. The reason is that eigenvalues come into play for elliptic equations. Therefore, the methods used in [2], [3], do not directly extend to the elliptic case that was not yet discussed in literature.

The novelty of this paper is twofold: First, the detour on linear programming is avoided by direct use of quadratic dual programs. In this way, the Lagrange multipliers are obtained by a constructive approach that is fairly elementary. Second, and this is the main issue, an estimation technique is applied that avoids the use of inverse monotonicity. An assumption of smallness of G is not needed.

2. The dual problem. Let us first define the notion of Lagrange multipliers. We introduce the *Lagrange functional* $L = L(u, \mu, \lambda, \nu) : L^2(D)^4 \to \mathbb{R}$, by

$$L(u, \mu, \lambda, \nu) = f(u) + (u - G\,u - c\,,\,\mu) + (u - b\,,\,\lambda) - (u\,,\,\nu),$$

define for convenience the vector of Lagrange multipliers $\omega := (\mu, \lambda, \nu)$, and write $\omega \geq 0$ whenever $\mu(x) \geq 0$, $\lambda(x) \geq 0$ and $\nu(x) \geq 0$ holds almost everywhere on D. Moreover, $\mu \geq 0$ means $\mu(x) \geq 0$ almost everywhere on D.

Definition. *Let $\bar u$ be a solution of the problem (PP). Functions $\bar\mu$, $\bar\lambda$, $\bar\nu \in L^2(D)$ are said to be Lagrange multipliers associated with $\bar u$, if the following relations are satisfied:*

$$\frac{\partial L}{\partial u}(u, \bar\mu, \bar\lambda, \bar\nu) = 0$$

and

$$\begin{array}{ll} (\bar u - G\,\bar u - c\,,\,\bar\mu) = 0, & \bar\mu \geq 0 \\ (\bar u - b\,,\,\bar\lambda) = 0, & \bar\lambda \geq 0 \\ (\bar u\,,\,\bar\nu) = 0, & \bar\nu \geq 0. \end{array}$$

The second block of relations forms the well-known *complementary slackness conditions.* In convex optimization, Lagrange multipliers can be obtained as solutions of *dual problems*, [7]. There are different, mainly equivalent ways to establish them. In this paper, we invoke the concept of Lagrange duality.

It is easy to verify that the primal problem can be equivalently expressed by

$$(PP) \qquad \inf_{u \in L^2(D)} \{ \sup_{\omega \in L^2(D)^3,\, \omega \geq 0} L(u, \omega) \}.$$

The *Lagrange dual problem* is defined by reversing the order of *inf* and *sup*,

$$(DP) \qquad \sup_{\omega \in L^2(D)^3,\, \omega \geq 0} \{ \inf_{u \in L^2(D)} L(u, \omega) \}.$$

We shall prove that the Lagrange function has a saddle point, i.e. a pair $(\bar u, \bar\omega)$ such that

$$L(\bar u, \omega) \leq L(\bar u, \bar\omega) \leq L(u, \bar\omega) \tag{2.1}$$

holds for all $u \in L^2(D)$ and all $\omega \in L^2(D)^3$, $\omega \geq 0$. The property of $(\bar u, \bar\omega)$ to be a saddle point is equivalent to the statement that $\bar u$ is a solution of (PP) and $\bar\omega$ solves (DP). Let us denote the dual objective functional by g,

$$g(\mu, \lambda, \nu) = g(\omega) = \inf_{u \in L^2(D)} L(u, \omega).$$

In this way, the dual problem is equivalent to

$$(DP) \qquad \sup g(\omega), \quad \omega \geq 0,\ \omega \in L^2(D)^3.$$

Let us derive a more explicit expression for g.

LEMMA 2.1. *The dual objective function g is equal to*

$$g(\mu, \lambda, \nu) = -(c\,,\,\mu) - (b\,,\,\lambda) - \frac{1}{2}\|a + \mu - G^\star\mu + \lambda - \nu\|_\Lambda^2,$$

where the norm $\|\cdot\|_\Lambda$ *is defined by* $\Lambda = \kappa I + S^\star S$ *and*

$$\|d\|_\Lambda^2 = (d\,,\,\Lambda^{-1}d)\,.$$

Proof: The result is obtained by straightforward computations. We get

$$\begin{aligned} g(\mu,\lambda,\nu) = \min_{u\in L^2(D)} &L(u,\mu,\lambda,\nu) = -(c\,,\,\mu) - (b\,,\,\lambda) \\ &+ \min_u \{f(u) + (u - G\,u\,,\,\mu) + (u\,,\,\lambda) - (u\,,\,\nu)\}. \end{aligned} \tag{2.2}$$

The function in braces is convex differentiable and attains its unique minimum at $\hat{u} \in L^2(D)$. We find $\hat{u}$ by setting the derivative equal to zero,

$$\kappa\,\hat{u} + S^\star S\hat{u} + a + \mu - G^\star\mu + \lambda - \nu = 0,$$

hence

$$\hat{u} = -\Lambda^{-1}(a + \mu - G^\star\mu + \lambda - \nu) = -\Lambda^{-1}d, \tag{2.3}$$

where

$$d = a + \mu - G^\star\mu + \lambda - \nu. \tag{2.4}$$

Inserting $\hat{u}$ defined by (2.3) in $f(u) + (u - G\,u\,,\,\mu) + (u\,,\,\lambda) - (u\,,\,\nu)$, the formula for g is found instantly,

$$\begin{aligned} f(\hat{u}) + (\hat{u} - G\,\hat{u}\,,\,\mu) + (\hat{u}\,,\,\lambda - \nu) &= \frac{1}{2}\,\|S\hat{u}\|_H^2 + \frac{\kappa}{2}\,\|\hat{u}\|^2 + (\hat{u}\,,\,a + \mu - G^\star\mu + \lambda - \nu) \\ &= \frac{1}{2}\,(S\Lambda^{-1}d\,,\,S\Lambda^{-1}d)_H + \frac{\kappa}{2}\,(\Lambda^{-1}d\,,\,\Lambda^{-1}d) \\ &\quad -(d\,,\,\Lambda^{-1}d) \\ &= (\frac{1}{2}\,[\kappa\,\Lambda^{-1}d + S^\star S\Lambda^{-1}d] - d\,,\,\Lambda^{-1}d) \\ &= (\frac{1}{2}\,\Lambda\,\Lambda^{-1}d - d\,,\,\Lambda^{-1}d) = -\frac{1}{2}(d\,,\,\Lambda^{-1}d) \\ &= -\frac{1}{2}\,\|d\|_\Lambda^2. \end{aligned}$$

Now the formula for g is obtained from (2.2). ∎

LEMMA 2.2. *Assume the functions* b *and* c *to be non-negative. Then the primal problem (PP) admits a unique optimal solution* $\bar{u}$.

Proof: The arguments are standard. The non-negativity of b and c ensures that the feasible set of (PP) is not empty, since $u = 0$ satisfies all constraints. The objective functional f is strictly convex, continuous and tends to infinity as $\|u\| \to \infty$. Therefore the search for the minimum can be restricted to a convex and bounded set that is weakly sequentially compact. A weakly converging subsequence $\{u_n\}$ can be selected that tends to the optimal solution. We omit the standard details. Notice that the feasible set of (PP) is convex and closed, due to the continuity of G. ∎

Let us denote for convenience the negative objective functional $-g$ of (DP) by φ,

$$\varphi(\omega) = (c\,,\,\mu) + (b\,,\,\lambda) + \frac{1}{2}\|a + \mu - G^\star\mu + \lambda - \nu\|_\Lambda^2.$$

It is obvious that the dual problem (DP) is equivalent to minimizing φ.

LEMMA 2.3 (Weak duality). *For all feasible $u \in L^2(D)$ and all non-negative $\omega \in L^2(D)^3$, it holds $g(\omega) \le f(u)$.*

Proof: We have

$$\begin{aligned} f(u) &\ge f(u) + (u - G\,u - c\,,\,\mu) + (u - b\,,\,\lambda) - (u\,,\,\nu) \\ &= -(\mu\,,\,c) - (\lambda\,,\,b) + \{\frac{1}{2}\,\|Su\|_H^2 + \frac{\kappa}{2}\,\|u\|^2 + (u\,,\,a + \mu - G^\star\mu + \lambda - \nu)\}, \end{aligned}$$

since u is feasible for (PP) and μ, λ, ν are non-negative. Hence

$$\begin{aligned} f(u) &\ge -(\mu\,,\,c) - (\lambda\,,\,b) + \min_v \{\frac{1}{2}\,\|Sv\|^2 + \frac{\kappa}{2}\,\|v\|^2 + (v\,,\,a + \mu - G^\star\mu + \lambda - \nu)\} \\ &= g(\omega) \end{aligned}$$

by (2.2). ∎

Under natural assumptions, separation arguments apply to show that the optimal values of the primal and dual problem coincide, i.e., that there is no duality gap. To prove this, we need the following convex set being an epigraph associated with the primal problem,

$$\begin{aligned} K := \{(r,z,v,w) \in \mathbb{R} \times L^2(D)^3 \mid\ & \exists u \in L^2(D)\ : \\ & f(u) \le r,\ u - Gu - c \le z,\ u - b \le v,\ -u \le w\}. \end{aligned}$$

LEMMA 2.4. *The set K is convex and closed.*

Proof: Let $\{(r_n, z_n, v_n, w_n)\}$ be a sequence of elements of K converging to (r,z,v,w). We have to show that $(r,z,v,w) \in K$. By definition of K, functions u_n exist such that

$$f(u_n) \le r_n, \quad u_n - G\,u_n - c \le z_n, \quad u_n - b \le v_n, \quad -u_n \le w_n.$$

Since $r_n \ge f(u_n) \ge \kappa\,\|u_n\|^2$ and $r_n \to r$, $\{u_n\}$ must be bounded in $L^2(D)$ so that we can select a weakly converging subsequence, say $u_n \rightharpoonup u$. Thanks to lower semicontinuity of f, it holds $f(u) \le r$. Moreover, the set

$$\{(u,z,v,w) \in L^2(D)^4 \,|\, u - G\,u - c \le z,\ u - b - v \le 0,\ -u \le w\}$$

is convex and closed, hence weakly closed. Therefore, the limit (u,z,v,w) belongs to this set. Together with the result above, $(r,v,z,w) \in K$ is shown. ∎

THEOREM 2.5 (Duality). *Assume that $c(x) > 0$ and $b(x) > 0$ holds almost everywhere on D. Then the duality relation* inf *(PP)* = sup *(DP), i.e.*

$$f(\bar u) = \sup_{\omega \ge 0} g(\omega). \tag{2.5}$$

is satisfied.

Proof: Let $\bar u$ be the optimal solution of (PP). We take an arbitrary but fixed natural number n. Then the point $(f(\bar u) - 1/n, 0, 0, 0)$ does not belong to K. Notice that for each $u \in L^2(D)$ and all non-negative r, z, v, w the element $(\tilde r = f(u) + r, \tilde z = u - G\,u - c + z, \tilde v = u - b + v, \tilde w = -u + w)$ belongs to K. Since K is closed, we can separate this point from K by a separating hyperplane, i.e. there is an element $(\sigma, \mu, \lambda, \nu) \neq 0$ such that

$$\begin{aligned} \sigma\,(f(\bar u) - 1/n) + 0 \le \sigma\,(f(u) + r) + (\mu\,,\, u - G\,u - c + z) \\ +(\lambda\,,\, u - b + v) + (\nu\,,\, -u + w) \end{aligned} \tag{2.6}$$

for all u and all $r \ge 0$, $z \ge 0$, $v \ge 0$, $w \ge 0$. It is obvious that this relation can only hold, if σ, μ, λ, and ν are non-negative. Next we show $\sigma \neq 0$. Assume the contrary. Since b and c are almost everywhere positive, the function $\tilde u(x) = 0.5\,\min\{b(x), c(x)\}$ is almost everywhere positive and less than $\min\{b(x), c(x)\}$. By $G\,\tilde u \ge 0$ we have

$$\tilde u(x) - G\,\tilde u(x) - c(x) < 0 \quad \text{and} \quad 0 < \tilde u(x) < b(x) \quad \text{a.e. on } E.$$

In this sense, $\tilde u$ is a "weak" Slater point. Inserting $\tilde u$ in (2.6), we would get in view of the temporary assumption $\sigma = 0$

$$0 \le (\mu\,,\, \tilde u - G\,\tilde u - c) + (\lambda\,,\, \tilde u - b) - (\nu\,,\, \tilde u).$$

The right-hand side is strictly negative unless $\mu = \lambda = \nu = 0$, contradicting the property $(\sigma, \mu, \lambda, \nu) \neq 0$. Therefore we can assume $\sigma > 0$, without limitation of generality $\sigma = 1$ (otherwise we can divide the inequality by σ). Now we have

$$f(\bar u) - 1/n \le (f(u) + r) + (\mu\,,\, u - G\,u - c + z) + (\lambda\,,\, u - b + v) + (\nu\,,\, -u + w)$$

for all suitable u, r, v, z, w. We insert $r = 0$, $z = v = w = 0$. Then

$$\begin{aligned} f(\bar u) - 1/n &\le -(\mu\,,\, c) - (\lambda\,,\, b) + (\mu\,,\, u - G\,u) + (\lambda\,,\, u) - (\nu\,,\, u) + (a, u) \\ &\quad + \frac{1}{2}\,\|Su\|_H^2 + \frac{\kappa}{2}\,\|u\|^2 \\ &= -(\mu\,,\, c) - (\lambda\,,\, b) + (\mu - G^\star\mu + a + \lambda - \nu\,,\, u) + \frac{1}{2}\,\|Su\|_H^2 + \frac{\kappa}{2}\,\|u\|^2. \end{aligned}$$

Next we insert the function $\hat u = -\Lambda^{-1} d$ defined in (2.3). Then

$$\begin{aligned} f(\bar u) - 1/n &\le -(\mu\,,\, c) - (\lambda\,,\, b) - (d\,,\, \Lambda^{-1}d) + \frac{\kappa}{2}\,\|\Lambda^{-1}d\|^2 + \frac{1}{2}\,\|S\Lambda^{-1}d\|_H^2 \\ &= -(\mu\,,\, c) - (\lambda\,,\, b) - (d\,,\, \Lambda^{-1}d) + \frac{\kappa}{2}\,(\Lambda^{-1}d\,,\, \Lambda^{-1}d) \\ &\quad + \frac{1}{2}\,(S\Lambda^{-1}d\,,\, S\Lambda^{-1}d)_H \\ &= -(\mu\,,\, c) - (\lambda\,,\, b) - \frac{1}{2}(d\,,\, \Lambda^{-1}d) = g(\mu, \lambda, \nu) = g(\omega). \end{aligned}$$

In view of Lemma 2.3, it holds $g(\omega) \le f(\bar u)$. Since n was arbitrary, we have found an $\omega = \omega_n = (\mu_n, \lambda_n, \nu_n)$ such that

$$f(\bar u) - \frac{1}{n} \le g(\omega_n) \le f(\bar u).$$

The claim of the Lemma is proven for $n \to \infty$. ∎

The next results are standard. Their proofs are included for convenience only.

Lemma 2.6. *Let $\bar u$ be the solution of (PP) and $\bar\omega = (\bar\mu, \bar\lambda, \bar\nu) \ge 0$ be a solution of the dual problem (DP). Then the complementary slackness conditions*

$$(\bar u - G\,\bar u - c\,,\, \bar\mu) = (\bar u - b\,,\, \bar\lambda) = (\bar u\,,\, \bar\nu) = 0$$

are satisfied and the pair $\{\bar u, \bar\omega\}$ is a saddle point of the Lagrange functional L.

Proof: We know $f(\bar u) = g(\bar\omega)$ from Theorem 2.5, hence

$$
\begin{aligned}
f(\bar u) = g(\bar\omega) &= \min_{u \in L^2(D)} \{ f(u) + (u - Gu - c\,,\, \bar\mu) + (u - b\,,\, \bar\lambda) - (u\,,\, \bar\nu)\} \\
&\le f(\bar u) + (\bar u - G\bar u - c\,,\, \bar\mu) + (\bar u - b\,,\, \bar\lambda) - (\bar u\,,\, \bar\nu) \le f(\bar u) \qquad (2.7)
\end{aligned}
$$

since $\bar u$ is feasible and $\bar\mu$, $\bar\lambda$, $\bar\nu$ are non-negative. Therefore,

$$(\bar u - G\,\bar u - c\,,\, \bar\mu) + (\bar u - b\,,\, \bar\lambda) - (\bar u\,,\, \bar\nu) = 0$$

must hold. All scalar products above are non-positive, hence they vanish alltogether. The right-hand side of (2.7) is the value $L(\bar u, \bar\omega)$, hence

$$f(\bar u) = L(\bar u, \bar\omega) = g(\bar\omega) = \min_v \, L(v, \bar\omega) \le L(u, \bar\omega)$$

for all $u \in L^2(D)$. This is one half of the saddle point property. On the other hand,

$$L(\bar u, \omega) = f(\bar u) + (\bar u - G\bar u - c\,,\, \mu) + (\bar u - b\,,\, \lambda) - (\bar u\,,\, \nu) \le f(\bar u) = L(\bar u, \bar\omega)$$

because all scalar products are non-positive for feasible $\bar u$. This is the second part. ∎

THEOREM 2.7. *Suppose that a feasible function $\bar u$ is optimal for (PP) and the dual problem (DP) admits a solution $\bar\omega = (\bar\mu, \bar\lambda, \bar\nu)$. Then $\bar\omega$ is a triplet of Lagrange multipliers associated with $\bar u$.*

This is a simple consequence of the preceding Lemma. From the saddle point property we know that $\bar u$ solves the problem $\min\limits_u L(u, \bar\omega)$. Therefore, the derivative of L with respect to u must vanish at $\bar u$. Together with the complementary slackness conditions, this shows that $\bar\omega$ is a vector of Lagrange multipliers.

3. Existence of regular Lagrange multipliers. In this section, we prove the main result of this paper, the solvability of the dual problem. By Theorem 2.7, this implies the existence of Lagrange multipliers in the space $L^2(D)$. The proof is based upon the following assumption:

Assumption of non-negativity and boundedness. *The operator G is non-negative, i.e.*

$$u(x) \ge 0 \;\; a.e. \;\; on\; D \quad \Rightarrow \quad (Gu)(x) \ge 0 \;\; a.e. \;\; on\; D.$$

Moreover, there are some $r > 0$ and a constant $c_r > 0$ such that $G^\star$ is bounded from $L^1(D)$ to $L^{1+r}(D)$ and from $L^2(D)$ to $L^{2+r}(D)$, i.e.

$$\|G^\star \mu\|_{L^{i+r}(D)} \le c_r \, \|\mu\|_{L^i(D)} \qquad \forall\, \mu \in L^i(D),\; i = 1, 2.$$

Remark. Let us define

$$s = \min\{1 + r, 2\},$$

then $G^\star$ is bounded from $L^s(D)$ to $L^s(D)$. Moreover, we obtain by interpolation that

$$\|G^\star \mu\|_{L^{p+r}(D)} \le c_r \, \|\mu\|_{L^p(D)} \qquad \forall\, \mu \in L^p(D),\; \forall p \in [1, 2].$$

The dual problem is equivalent to minimizing φ on the set of non-negative functions of $L^2(D)^3$,

$$\inf_{\omega\geq 0} (c\,,\,\mu) + (b\,,\,\lambda) + \frac{1}{2}\|a+\mu-G^\star\mu+\lambda-\nu\|_\Lambda^2. \tag{3.1}$$

Since $\varphi(\omega)\geq 0$ for all $\omega\geq 0$, the corresponding infimum j of φ exists as a non-negative real number. Let $\{\omega_n\}=\{(\mu_n,\lambda_n,\nu_n)\}$ be a minimizing sequence, i.e.

$$\lim_{n\to\infty}\varphi(\omega_n)=j.$$

Define d_n by

$$d_n = a+\mu_n-G^\star\mu_n+\lambda_n-\nu_n. \tag{3.2}$$

LEMMA 3.1. *Assume that $c(x)\geq\delta_0$ and $b(x)\geq\delta_0$ holds almost everywhere on D with some constant $\delta_0>0$. Then a constant $M>0$ exists such that*

$$\|d_n\|_{L^2(D)}+\|\mu_n\|_{L^1(D)}+\|\lambda_n\|_{L^1(D)}\leq M$$

holds for all natural n.

Proof: Notice that $b(x)\geq\delta_0$, $c(x)\geq\delta_0$ implies

$$(c\,,\,\mu)\geq\delta_0\,\|\mu\|_{L^1(D)},\quad (b\,,\,\lambda)\geq\delta_0\,\|\lambda\|_{L^1(D)}$$

for non-negative λ and μ. Hence, it holds for all sufficiently large n

$$\delta_0\,\|\mu_n\|_{L^1(D)}+\delta_0\,\|\lambda_n\|_{L^1(D)}+\frac{1}{2}\,\|d_n\|_\Lambda^2\leq j+1, \tag{3.3}$$

because $\{\omega_n\}$ is a minimizing sequence. Otherwise $\varphi(\omega_n)$ would tend to infinity. Therefore, the L^1-norms of μ_n and λ_n are bounded. The operator $\Lambda=\kappa\,I+S^\star S$ is positive definite, as

$$(e\,,\,\Lambda e)\geq\kappa\,\|e\|^2$$

holds for all $e\in L^2(D)$. This implies that, for all $d\in L^2(D)$,

$$\|d\|_\Lambda^2=(d\,,\,\Lambda^{-1}d)=(\Lambda\,(\Lambda^{-1}d)\,,\,\Lambda^{-1}d)\geq\kappa\,\|\Lambda^{-1}d\|^2\geq\kappa\,c_\Lambda^{-2}\|d\|^2.$$

Here, we have used the boundedness of Λ, $\|\Lambda\,e\|\leq c_\Lambda\,\|e\|$, hence $\|d\|\leq c_\Lambda\,\|\Lambda^{-1}d\|$.

In view of this, unboundedness of $\|d_n\|$ would imply unboundedness of $\|d_n\|_\Lambda$, contradicting (3.3). Therefore, $\|d_n\|$ is bounded, too. ∎

LEMMA 3.2. *Let the assumptions of the preceding Lemma 3.1 be satisfied and let $\omega=(\mu,\lambda,\nu)\geq 0$ be given. Then there is another $\tilde\omega=(\tilde\mu,\tilde\lambda,\tilde\nu)\geq 0$ and a constant c_b that does not depend on ω and $\tilde\omega$, such that*

$$\varphi(\tilde\omega)\leq\varphi(\omega) \tag{3.4}$$

and

$$\tilde\mu(x)+\tilde\lambda(x)+\tilde\nu(x)\leq 2\,(\,|d(x)|+|a(x)|+(G^\star(\mu+\lambda))(x)\,). \tag{3.5}$$

holds almost everywhere on D.

Proof: We define the measurable sets

$$\begin{aligned} I_0 &= \{x \in E \,|\, \nu(x) = 0\} \\ I_+ &= \{x \in E \,|\, \nu(x) > 0\}. \end{aligned}$$

(i) *Update of* (μ, λ)

We update μ and λ by $\tilde\mu + \tilde\lambda = \mu + \lambda - \delta$ with a certain function $\delta = \delta(x)$ to defined below.

On I_0 we have by (2.4)

$$\mu(x) + \lambda(x) = d(x) - a(x) + (G^\star\, \mu)(x).$$

The sum $\mu + \lambda$ is already bounded on I_0 by an expression of the type (3.5). There is no reason to update μ and λ on I_0, and we define the update $\delta(x)$ by

$$\delta(x) = 0 \text{ on } I_0.$$

On I_+ we consider the two sets

$$\begin{aligned} I_+(\nu) &= \{x \in I_+ \,|\, \mu(x) + \lambda(x) \ge \nu(x)\} \\ I_+(\mu) &= \{x \in I_+ \,|\, \mu(x) + \lambda(x) < \nu(x)\} \end{aligned}$$

and put

$$\delta(x) = \begin{cases} \nu(x) & \text{for } x \in I_+(\nu) \\ \mu(x) + \lambda(x) & \text{for } x \in I_+(\mu). \end{cases}$$

Now, δ is defined on the whole set D, and we update μ and λ by

$$\tilde\mu(x) + \tilde\lambda(x) = \mu(x) + \lambda(x) - \delta(x)$$

while ensuring $0 \le \tilde\mu(x) \le \mu(x)$ and $0 \le \tilde\lambda(x) \le \lambda(x)$ for all $x \in E$. This can be accomplished in many ways. To fix the setting, let us make the choice $\delta = \delta_\mu + \delta_\lambda$, where

$$\delta_\mu(x) = \begin{cases} 0 & \text{for } x \in I_0 \\ \nu(x)/2 & \text{for } x \in I_+(\nu) \\ \mu(x) & \text{for } x \in I_+(\mu) \end{cases} \qquad \delta_\lambda(x) = \begin{cases} 0 & \text{for } x \in I_0 \\ \nu(x)/2 & \text{for } x \in I_+(\nu) \\ \lambda(x) & \text{for } x \in I_+(\mu) \end{cases}$$

and $\tilde\mu = \mu - \delta_\mu$, $\tilde\lambda = \lambda - \delta_\lambda$.

(ii) *Bound for* $(\tilde\mu, \tilde\lambda)$

In view of our definitions, it holds

$$\tilde\mu(x) + \tilde\lambda(x) = \begin{cases} \mu(x) + \lambda(x) = d(x) - a(x) + (G^\star\mu)(x) & \text{on } I_0 \\ \mu(x) + \lambda(x) - \nu(x) = d(x) - a(x) + (G^\star\mu)(x) & \text{on } I_+(\nu) \\ 0 & \text{on } I_+(\mu). \end{cases} \tag{3.6}$$

hence (3.6) implies in turn that

$$\tilde\mu(x) + \tilde\lambda(x) \le |d(x)| + |a(x)| + (G^\star\,(\mu + \lambda))(x). \tag{3.7}$$

holds a.e. on D. Here we have used $G^\star\,\mu \le G^\star\,(\mu+\lambda)$.

(iii) *Update of* ν

Next we adapt ν in a way leaving $\|d\|_\Lambda$ unchanged. By (2.4), it holds on D

$$\mu + \lambda = d - a + G^\star \mu + \nu.$$

Inserting $\mu+\lambda = \tilde\mu + \tilde\lambda + \delta$ and $\mu = \tilde\mu + \delta_\mu$, we find

$$\tilde\mu + \tilde\lambda = d - a + G^\star\tilde\mu + G^\star\delta_\mu + \nu - \delta.$$

Therefore, it is natural to define

$$\tilde\nu = G^\star\delta_\mu + \nu - \delta. \tag{3.8}$$

Then $\tilde\nu$ is non-negative on D, since $\nu - \delta \ge 0$, $\delta_\mu \ge 0$ and hence $G^\star\delta_\mu \ge 0$. Moreover,

$$\tilde\mu + \tilde\lambda + a - G^\star\tilde\mu - \tilde\nu = d \tag{3.9}$$

is satisfied everywhere on D.

On $I_0 \cup I_+(\nu)$, the relation $\nu(x) = \delta(x)$ holds true, hence (3.8) yields

$$\tilde\nu(x) = (G^\star\delta_\mu)(x) \quad \forall x \in I_0 \cup I_+(\nu). \tag{3.10}$$

Moreover, we know $\tilde\mu(x) + \tilde\lambda(x) = 0$ on $I_+(\mu)$ and obtain from (3.9) that

$$\tilde\nu(x) = a(x) - d(x) - (G^\star\tilde\mu)(x) \qquad \text{on } I_+(\mu).$$

Summarizing up, $\tilde\nu$ is defined by

$$\tilde\nu(x) = \begin{cases} (G^\star\delta_\mu)(x) \le (G^\star\delta)(x) \le (G^\star(\mu+\lambda))(x) & \text{on } I_+(\nu)\cup I_0 \\ a(x) - d(x) - (G^\star\tilde\mu)(x) & \text{on } I_+(\mu) \end{cases} \tag{3.11}$$

implying in turn

$$\tilde\nu(x) \le |d(x)| + |a(x)| + (G^\star\,(\mu+\lambda))(x)$$

a.e. on D, as $0 \le \tilde\mu \le \mu + \lambda$. Together with (3.7), this shows that $\tilde\omega$ satisfies (3.5).

(iv) *Verification of (3.4)*

By definition, we have

$$\begin{aligned} 0 &\le \tilde\mu(x) = \mu(x) - \delta_\mu(x) \le \mu(x) \\ 0 &\le \tilde\lambda(x) = \lambda(x) - \delta_\lambda(x) \le \lambda(x), \end{aligned}$$

hence

$$(c\,,\,\tilde\mu) \le (c\,,\,\mu) \qquad \text{and} \qquad (b\,,\,\tilde\lambda) \le (b\,,\,\lambda).$$

is satisfied. Moreover, our construction guarantees by (3.9) that

$$\tilde\mu + \tilde\lambda + a - G^\star\tilde\mu - \tilde\nu = d = \mu + \lambda + a - G^\star\mu - \nu.$$

This yields

$$\begin{aligned}\varphi(\tilde\omega) &= (c\,,\,\tilde\mu) + (b\,,\,\tilde\lambda) + \frac{1}{2}\,\|\tilde\mu + \tilde\lambda + a - G^\star\tilde\mu - \tilde\nu\|_\Lambda^2\\ &= (c\,,\,\tilde\mu) + (b\,,\,\tilde\lambda) + \frac{1}{2}\,\|\mu + \lambda + a - G^\star\mu - \nu\|_\Lambda^2\\ &\le (c\,,\,\mu) + (b\,,\,\lambda) + \frac{1}{2}\,\|\mu + \lambda + a - G^\star\mu - \nu\|_\Lambda^2 = \varphi(\omega).\end{aligned}$$

∎

Theorem 3.3. *Assume that G satisfies the assumption of non-negativity and boundedness, that b and c belong to $L^{s'}(D)$ with conjugate exponent s' defined by $1/s + 1/s' = 1$, and that $b(x) \ge \delta_0$ and $c(x) \ge \delta_0$ is fulfilled for almost all $x \in D$ with some constant $\delta_0 > 0$. Then the dual problem admits at least one solution.*

Proof: Let ω_n be the minimizing sequence used in Lemma 3.1. By this Lemma, the sequence $d_n = a + \mu_n - G^\star\,\mu_n + \lambda_n - \nu_n$ is bounded in $L^2(D)$ and we can assume w.l.o.g. that $d_n \rightharpoonup d$ in $L^2(D)$ to some $d \in L^2(D)$. In particular, this sequence is weakly converging in $L^s(D)$. Moreover, we obtain from Lemma 3.1 the boundedness of $\{\mu_n\}$ in $L^1(D)$. The assumptions imposed on $G^\star$ imply that the sequence $\{G^\star\,\mu_n\}$ is bounded in $L^s(D)$.

Thanks to Lemma 3.2, the associated sequence $\tilde\omega_n$ is bounded in $L^s(D)$, too, and we can assume without limitation of generality that $\tilde\omega_n \rightharpoonup \omega = (\mu, \lambda, \nu)$ in $L^s(D)$. The operator $G^\star$ is bounded in the L^s-norm, hence G is bounded in $L^{s'}(D)$, and we get the weak convergence of $\{G^\star\,\tilde\mu_n\}$ to $\{G^\star\,\mu\}$ in $L^s(D)$. Altogether, we have

$$d_n = a + \tilde\mu_n - G^\star\,\tilde\mu_n + \tilde\lambda_n - \tilde\nu_n,$$

where all sequences converge weakly in $L^s(D)$. Passing to the limit, we find

$$d = a + \mu - G^\star\,\mu + \lambda - \nu.$$

It is obvious that

$$(c\,,\,\tilde\mu_n) \to (c\,,\,\mu) \quad \text{ and } (b\,,\,\tilde\lambda_n) \to (b\,,\,\lambda)$$

as $n \to \infty$, because b and c are assumed to be in $L^{s'}(D)$. Moreover, the functional $\|\cdot\|_\Lambda^2$ is continuous and convex in $L^2(D)$, hence weakly lower semicontinuous. This yields

$$\begin{aligned} j &= \lim_{n\to\infty}\varphi(\omega_n) \ge \lim_{n\to\infty}\varphi(\tilde\omega_n) = \lim_{n\to\infty}\{(c\,,\,\tilde\mu_n) + (b\,,\,\tilde\lambda_n) + \|d_n\|_\Lambda^2\}\\ &= (c\,,\,\mu) + (b\,,\,\lambda) + \lim_{n\to\infty}\|d_n\|_\Lambda^2\\ &\ge (c\,,\,\mu) + (b\,,\,\lambda) + \|d\|_\Lambda^2\\ &= (c\,,\,\mu) + (b\,,\,\lambda) + \|a + \mu - G^\star\,\mu + \lambda - \nu\|_\Lambda^2 = \varphi(\omega),\end{aligned}$$

thus $\varphi(\omega) = j$.

It remains to show that the solution ω belongs to $L^2(D)$. This is done by a bootstrapping argument. So far, we have $\omega \in L^s(D)$ and

$$0 \le \mu(x) + \lambda(x) + \nu(x) \le 2\,(|d(x)| + |a(x)| + (G^\star\,(\mu + \lambda))(x)).$$

We know that d and a belong to $L^2(D)$ and obtain $G^\star\,(\mu + \lambda) \in L^{s+r}(D)$ by the smoothing property of $G^\star$. In this way, the inequality above delivers the regularity

$\omega \in L^{\min\{s+r,2\}}(D)$. Repeating this argument, in the next step $\omega \in L^{\min\{s+2r,2\}}(D)$ is obtained. After finitely many steps we arrive at the regularity $\omega \in L^2(D)$. ∎

Remark. If the property $\|d_1\| \leq \|d_2\| \Rightarrow \|d_1\|_\Lambda \leq \|d_2\|_\Lambda$ holds true, then higher regularity of ω might be proven, provided that a, b, c are more regular. For instance, L^∞-regularity of multipliers can be expected. We do not discuss this point here.

4. Other types of constraints. Let us mention a few more interesting cases, where the theory applies after minor changes.

4.1. Single control-state constraint. Consider the optimization problem

$$(P1) \qquad \min f(u), \quad u(x) \le c(x) + (G\,u)(x) \quad \text{a.e. on D}$$

without any box constraint on u. Existence and uniqueness of an optimal solution $\bar{u}$ follows from the strict convexity of f. The existence of an associated regular Lagrange multiplier $\bar{\mu}$ can be proven in a direct and rather trivial way. We assume that $I - G$ has a continuous inverse operator in $L^2(D)$. In the applications to PDEs, G is compact, hence this amounts to the assumption that $\alpha = 1$ is not an eigenvalue of G.

We substitute $v = u - G\,u$ and have $u = (I - G)^{-1} v = R\,v$, if we set $R = (I - G)^{-1}$. Define $\tilde{f}(v) = f(R\,v)$. Then problem (P1) is equivalent to

$$(\tilde{P}1) \qquad \min \tilde{f}(v), \quad v(x) \le c(x) \quad \text{a.e. in } D,$$

where $\tilde{f}(v) = f(R\,v)$. This is a problem without state constraints. For the optimal solution $\bar{v} = (I - G)\,\bar{u}$ we obtain by standard methods the variational inequality

$$\tilde{f}'(\bar{v})(v - \bar{v}) \ge 0 \quad \forall v \le c. \tag{4.1}$$

By the Riesz representation theorem, the gradient $f'(\bar{v})$ can be identified with a function of $L^2(D)$. Then (4.1) implies a.e. on D

$$\tilde{f}'(\bar{v})(x) \begin{cases} = 0 & \text{where } \bar{v}(x) \le c(x) \\ \le 0 & \text{where } \bar{v}(x) = c(x). \end{cases}$$

We just *define* the multiplier $\bar{\mu}$ by

$$\bar{\mu}(x) := -\tilde{f}'(\bar{v})(x). \tag{4.2}$$

This definition assures $\bar{\mu} \ge 0$,

$$\tilde{f}'(\bar{v}) + \bar{\mu} = 0 \tag{4.3}$$

and

$$(\bar{v} - c\,,\, \bar{\mu}) = 0. \tag{4.4}$$

THEOREM 4.1. *Assume that $(I - G)^{-1}$ exists and is continuous in $L^2(D)$. Then the function $\bar{\mu}$ defined in (4.2) is a Lagrange multiplier associated with the optimal solution $\bar{u}$ of (P1).*

Proof. The function $\bar{u}$ is optimal for (P1) if and only if $\bar{v} = \bar{u} - G\,\bar{u}$ is optimal for ($\tilde{\text{P}}$1). By definition, $\bar{\mu}$ is non-negative. Moreover, inserting the concrete expression for $\bar{v}$, (4.4) yields

$$(\bar{u} - G\,\bar{u} - c\,,\, \bar{\mu}) = 0,$$

and the complementary slackness condition is satisfied. We have

$$\tilde{f}'(\bar{v}) = R^\star f'(R\,\bar{v}) = R^\star f'(\bar{u})$$

and $L(u,\mu) = f(u) + (u - Gu - c\,,\,\mu)$. Therefore, $\bar\mu$ is a Lagrange multiplier, provided that also $\partial L/\partial u = 0$ holds, i.e.

$$f'(\bar u) + (I - G^\star)\mu = 0. \tag{4.5}$$

In view of $\tilde f'(\bar v) = R^\star f'(\bar u)$, (4.3) implies

$$R^\star f'(\bar u) + \bar\mu = 0,$$

which is equivalent to (4.5). ∎

4.2. The case $b = \infty$. Consider the constraints

$$0 \le u(x) \le c(x) + (Gu)(x).$$

Here, the upper bound b is not imposed on u. This case fits into our setting by the formal choice $b = \infty$. The reader will easily verify that Theorem 3.3 remains true with $\bar\lambda = 0$. The assumption $b(x) \ge \delta_0 > 0$ is not needed, it is formally satisfied.

4.3. Arbitrary box constraints. Let the more general box constraints

$$a(x) \le u(x) \le b(x), \qquad u(x) \le c(x) + (Gu)(x)$$

be given. In this case, we set $v = u - a$. Then the constraints transform to

$$0 \le v \le b - a, \qquad v \le c - a + G\,a + G\,v.$$

The theory applies after introducing $\tilde c = c - a + G\,a$. Here, the strict positivity of c required in Theorem 3.3 means $a + \delta_0 \le c + G\,a$. In other words, the lower bound a must strictly satisfy the state-constraints. Then a triplet of Lagrange multipliers exists by Theorem 3.3.

4.4. Different orientation of constraints. Another interesting type of constraints has the form

$$a(x) \le u(x) \le b(x), \qquad c(x) + (Gu)(x) \le u(x).$$

These constraints can be transformed to our initial setting as well. We put $v = b - u$. Then the new constraints are

$$0 \le v \le b - a, \qquad v \le b - c - G\,b + G\,v.$$

We substitute $\tilde c = b - c - G\,b$, and the strict positivity of $\tilde c$ amounts to the requirement that the upper bound b strictly fulfills the state constraint. If this assumption is satisfied, then Theorem 3.3 yields the existence of regular Lagrange multipliers.

4.5. Twosided control-state constraints. Constraints of the type

$$u(x) \le b(x), \qquad c_1(x) + (Gu)(x) \le u(x) \le c_2(x) + (Gu)(x)$$

can be reduced to the case discussed in subsection 4.4 by the transformation $v = u - G\,u$ that already has been applied in 4.1. We leave this to the reader. The same holds for $a \le u$ instead of $u \le b$. However, we were not able to prove the regularity of Lagrange multipliers, if the control constraints and the mixed control-state constraints are simultaneously twosided.

4.6. Negative sign with the control. For several reasons, the discussion of the constraints

$$u \geq 0, \qquad -u \leq c + G\,u.$$

is useful, too. Here, we cannot require $c \geq 0$, because otherwise $-u \leq 0 \leq c + G\,u$ would be automatically satisfied for all non-negative u.

The negative sign in front of the control cannot be bypassed by a transformation to the standard case. However, this case turns out to be even simpler. The proof of existence of $\bar{\mu}$ in $L^2(D)$ is rather easy. The only point, where the the negative sign influences the theory appears in (3.1). We have to show that the minimization in (3.1) admits a solution. To this aim, let us consider the associated problem (3.1). Now, it reads

$$\min_{\mu \geq 0,\, \nu \geq 0} (c\,,\,\mu) + \frac{1}{2}\|a - \mu - G^\star \mu - \nu\|_\Lambda^2.$$

We show that this problem admits a solution. Let $\{\mu_n, \nu_n\}$ be an associated minimizing sequence. Then $\|a - \mu_n - G^\star \mu_n - \nu_n\|$ must be bounded, since otherwise $\|a - \mu_n - G^\star \mu_n - \nu_n\|_\Lambda$ would tend to infinity. Therefore, $\|\mu_n + G^\star \mu_n + \nu_n\|$ is bounded as well. All functions under this norm are nonnegative, hence $\|\mu_n\| + \|\nu_n\|$ are bounded, too. This is the decisive point, where we can select weakly converging subsequences, w.l.o.g. $\mu_n \to \bar{\mu}$ and $\nu_n \to \bar{\nu}$. The optimality of $\bar{\mu}$, $\bar{\nu}$ follows by lower semicontinuity of the objective functional f. The remaining part of the theory is along the lines of the preceding section. In this way, we have proven the following result:

THEOREM 4.2. *Assume that G satisfies the assumption of non-negativity and boundedness. Then the problem*

$$(P2) \qquad \min f(u), \qquad u \geq 0, \quad -u \leq c + G\,u$$

has a unique solution $\bar{u}$. Moreover, associated regular Lagrange multipliers $\bar{\mu}$, $\bar{\nu}$ exist in $L^2(D)$.

The constraints $u \leq 0, \quad G\,u + c \leq -u$ can be transformed to the form of the constraints in (P2) by substituting $v = -u$.

5. Application to the examples (i) – (iii).

5.1. Elliptic distributed problem (i). In this problem, S was the solution operator for the elliptic boundary value problem, $G = S$, $a = -S^\star\, y_\Omega$. Moreover, it is known that the operator $S^\star z$ is given by $S^\star z = p$, where $p \in H_0^1(\Omega)$ solves the adjoint boundary value problem

$$\begin{aligned} -\Delta\, p &= z \\ p|_\Gamma &= 0. \end{aligned}$$

From $\partial L/\partial u = 0$ it follows

$$S^\star(S\bar{u} - y_\Omega - \bar{\mu}) + \kappa\,\bar{u} + \bar{\mu} + \bar{\lambda} - \bar{\nu} = 0. \tag{5.1}$$

Consequently, we introduce an adjoint state p by $p = S^\star(\bar{y} - y_\Omega - \bar{\mu})$,

$$\begin{aligned} -\Delta\, p &= \bar{y} - y_\Omega - \bar{\mu} \\ p|_\Gamma &= 0. \end{aligned}$$

Then (5.1) reads $p + \kappa\,\bar u + \bar\mu + \bar\lambda - \bar\nu = 0$. The control $u = \bar u$ with state $y = \bar y = S\,\bar u$ is optimal, if and only of the optimality system

$$\begin{aligned} -\Delta\, y &= u & \qquad -\Delta\, p &= y - y_\Omega - \mu \\ y|_\Gamma &= 0 & p|_\Gamma &= 0 \end{aligned}$$

$$p + \kappa\, u + \mu + \lambda - \nu = 0$$

$$\begin{array}{c} u(x) \ge 0, \quad \nu(x) \ge 0, \quad u(x)\nu(x) = 0 \\ u(x) \le b(x), \quad \lambda(x) \ge 0, \quad (u(x) - b(x))\,\lambda(x) = 0 \\ u(x) - c(x) - y(x) \le 0, \quad \mu(x) \ge 0, \quad (u(x) - c(x) - y(x))\,\mu(x) = 0 \end{array}$$

is satisfied for almost all $x \in \Omega$ with the corresponding adjoint state p and associated Lagrange multipliers μ, λ, ν from $L^2(\Omega)$. In the system above, $(\cdot\,,\,\cdot)$ stands for the inner product of $L^2(\Omega)$.

The existence of the Lagrange multipliers follows from Theorem 3.3, because the operator G satisfies all assumptions. Its non-negativity follows from the maximum principle for elliptic equations. Moreover, $G = G^\star$ is bounded from $L^1(\Omega) \to L^s(\Omega)$ for all $s < N/(N-2)$ [5] and from $L^2(\Omega)$ to $H^1_0(\Omega) \subset L^s(\Omega)$ for all $s < 2N/(N-2)$. This assures the boundedness property.

Remark. The same system can formally be obtained by the Lagrange function as follows: We append all inequality constraints for u by Lagrange multipliers and consider the elliptic optimal control problem without inequality constraints:

Minimize

$$L(y,u,\mu,\lambda,\nu) = \frac{1}{2}\,\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\kappa}{2}\,\|u\|^2_{L^2(\Omega)} + (u - y - c\,,\,\mu) + (u - b\,,\,\lambda) - (u\,,\,\nu)$$

subject to

$$\begin{aligned} -\Delta\, y &= u \\ y|_\Gamma &= 0. \end{aligned}$$

In other words, we eliminate the pointwise constraints by Lagrange multipliers while keeping the PDE as an explicit constraint. Then we establish the standard first order necessary optimality conditions for this elliptic control problem. In this way, we arrive at the optimality system established above. We illustrate this technique below.

5.2. Elliptic boundary control problem (ii). Let us establish the optimality system in the formal way explained above. For given multipliers μ, λ, ν, we consider the following elliptic optimal control problem:

Minimize

$$\begin{aligned} L(y,u,\mu,\lambda,\nu) = {} & \tfrac{1}{2}\,\|\alpha y - y_\Omega\|^2_{L^2(\Omega)} + \tfrac{\kappa}{2}\,\|u\|^2_{L^2(\Omega)} \\ & + (u - \gamma\, y - c\,,\,\mu) + (u - b\,,\,\lambda) - (u\,,\,\nu) \end{aligned}$$

subject to

$$\begin{aligned} -\Delta\, y &= 0 \\ \partial_n y + \beta\, y &= u. \end{aligned}$$

In this problem, $(\cdot\,,\,\cdot)$ stands for the inner product of $L^2(\Gamma)$. We know that the adjoint state is given by the equations $-\Delta\, p = D_y L|_\Omega$ and $\partial_n p + \beta\, p = D_y L|_\Gamma$, hence the adjoint equation is

$$\begin{aligned} -\Delta\, p &= \alpha(\alpha y - y_\Omega)\\ \partial_n p + \beta\, p &= -\gamma\mu. \end{aligned}$$

Moreover,

$$p + \kappa\, u + \mu + \lambda - \nu = 0$$

must hold. The existence of regular Lagrange multipliers in $L^2(\Gamma)$ follows from Theorem 3.3. The associated assumptions are satisfied: Again, $G \geq 0$ follows from the maximum principle, and $G^\star$ transforms $L^1(\Gamma)$ into $L^s(\Gamma)$ for all $s < (N-1)/(N-2)$ [5] and $L^2(\Gamma)$ to $H^{1/2}(\Gamma) \subset L^s(\Gamma)$ for all $s < 2(N-1)/(N-2)$. Again, the boundedness property is easy to verify.

The optimality system reads

$$\begin{aligned} -\Delta\, y &= 0 & -\Delta\, p &= \alpha\,(\alpha y - y_\Omega)\\ \partial_n y + \beta\, y &= u & \partial_n p + \beta\, p &= -\gamma\,\mu \end{aligned}$$

$$p(x) + \kappa\, u(x) + \mu(x) + \lambda(x) - \nu(x) = 0$$

$$\begin{gathered} u(x) \geq 0, \quad \nu(x) \geq 0, \quad u(x)\nu(x) = 0\\ u(x) \leq b(x), \quad \lambda(x) \geq 0, \quad (u(x) - b(x))\,\lambda(x) = 0 \end{gathered}$$

$$\begin{gathered} u(x) - c(x) - \gamma(x)\, y(x) \leq 0, \quad \mu(x) \geq 0,\\ (u(x) - c(x) - \gamma(x)\, y(x))\mu(x) = 0 \end{gathered}$$

for almost all $x \in \Gamma$.

5.3. Parabolic control problem (iii). The parabolic example is discussed analogously. Theorem 3.3 assures the existence of regular Lagrange multipliers. The associated assumptions are met. Also here, the non-negativity of G follows from the maximum principle. Moreover, $G^\star$ is bounded from $L^1(\Sigma) \to L^s(\Sigma)$ for all $s < (N+1)/N$, [8], Thm. 4.2. and from $L^2(\Gamma)$ to $L^s(\Sigma)$, for all $s < 2(N+1)/(N-1)$ [8], Thm. 4.1. The boundedness property is satisfied. Applying the same formal technique, the optimality system

$$\begin{aligned} y_t - \Delta\, y &= 0 & -p_t - \Delta\, p &= 0\\ \partial_n y + \beta\, y &= u & \partial_n p + \beta\, p &= -\gamma\,\mu\\ y(\cdot, 0) &= 0 & p(\cdot, T) &= \alpha\,(\alpha\, y(T) - y_\Omega) \end{aligned}$$

$$(p + \kappa\, u + \mu + \lambda - \nu)(x,t) = 0$$

$$\begin{gathered} u(x,t) \geq 0, \quad \nu(x,t) \geq 0, \quad u(x,t)\nu(x,t) = 0\\ u(x,t) \leq b(x,t), \quad \lambda(x,t) \geq 0, \quad (u(x,t) - b(x,t))\,\lambda(x,t) = 0 \end{gathered}$$

$$\begin{gathered} u(x,t) - c(x,t) - \gamma(x,t)\, y(x,t) \leq 0, \quad \mu(x,t) \geq 0,\\ (u(x,t) - c(x,t) - \gamma(x,t)\, y(x,t))\,\mu(x,t) = 0 \end{gathered}$$

must be satisfied for almost all $(x,t) \in \Sigma = \Gamma \times (0,T)$.

REFERENCES

[1] N. Arada and J. P. Raymond. Optimal control problems with mixed control-state constraints. *SIAM J. Control*, 39:1391–1407, 2000.

[2] M. Bergounioux and F. Tröltzsch. Optimal control of linear bottleneck problems. *ESAIM, Control, Optimisation and Calculus of Variations*, 3:235–250, 1998.

[3] M. Bergounioux and F. Tröltzsch. Optimal control of semilinear parabolic equations with state-constraints of bottleneck type. *ESAIM, Control, Optimisation and Calculus of Variations*, 4:595–608, 1999.

[4] E. Casas, J.-P. Raymond, and H. Zidani. Pontryagins principle for local solutions of control problems with mixed control-state constraints. *SIAM J. Control and Optimization*, 39:1182–1203, 2000.

[5] E. Casas, F. Tröltzsch, and A. Unger. Second order sufficient optimality conditions for some state-constrained control problems of semilinear elliptic equations. *SIAM J. Control and Optimization*, 38(5):1369–1391, 2000.

[6] N. Levinson. A class of continuous linear programming problems. *J. Math. Analysis Appl.*, 16:73–83, 1966.

[7] D.G. Luenberger. *Optimization by Vector Space Methods*. Wiley, New York, 1969.

[8] J.-P. Raymond and F. Tröltzsch. Second order sufficient optimality conditions for nonlinear parabolic control problems with state constraints. *Discrete and Continuous Dynamical Systems*, 6:431–450, 2000.

[9] A. Rösch and F. Tröltzsch. Sufficient second order optimality condititions for a parabolic optimal control problem with pointwise control-state constraints. *SIAM J. Control and Optimization*, 42:138–154, 2003.

[10] F. Tröltzsch. A minimum principle and a generalized bang-bang-principle for a distributed optimal control problem with constraints on the control and the state. *ZAMM*, 59:737–739, 1979.

[11] W.F. Tyndall. A duality theorem for a class of continuous linear programming problems. *J. SIAM*, 13:644–666, 1965.